Refine Search

Your wildcard search against 10000 terms has yielded the results below.

Your result set for the last L# is incomplete.

The probable cause is use of unlimited truncation. Revise your search strategy to use limited truncation.

Search Results -

Terms	Documents
(categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key-phrases or (key adj phrase$))) and (dynamic$3 or authomatic$3)and (select$3 same subset$)	2

Database:
US Pre-Grant Publication Full-Text Database
US Patents Full-Text Database
US OCR Full-Text Database
EPO Abstracts Database
JPO Abstracts Database
Derwent World Patents Index
IBM Technical Disclosure Bulletins

Search: L23

Refine Search | Recall Text | Clear | Interrupt

Search History

DATE: Tuesday, August 29, 2006 **Purge Queries** Printable Copy Create Case

Set Name side by side	Query	Hit Count	Set Name result set
	DB=PGPB,USPT,USOC,EPAB,JPAB,DWPI,TDBD; PLUR=YES; OP=ADJ		
L23	(categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key-phrases or (key adj phrase$))) and (dynamic$3 or authomatic$3)and (select$3 same subset$)	2	L23
L22	(categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key-phrases or (key adj phrase$))) and (dynamic$3 or authomatic$3)	2	L22
L21	L20 and @pd > 20060828	0	L21
L20	((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key-phrases or (key adj phrase$)))).ab.	0	L20

L19	709/$.ccls. and ((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key adj phrase$ or phrase$)))	0	L19
L18	709/$.ccls. and((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key adj phrase$ or "key-phrases")))and displays and (database near search$ or quer$)))	0	L18
L17	709/$.ccls. and((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same (key adj phrase$ or phrase$)))and displays and (database near search$ or quer$)))	0	L17
L16	L15 and @pd > 20050513	1	L16
L15	L14 and L13	7	L15
L14	709/$.ccls. and((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$)))	19	L14
L13	(707/1 |707/2 |707/3 |707/4 |707/5 |707/6 |707/7 |707/8 |707/9 |707/10 |707/100 |707/101 |707/102 |707/103R |707/103Y |707/103X |707/103Z |707/104.1 |707/200 |707/201 |707/202 |707/203 |707/204 |707/205 |707/206).ccls. and ((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$)))	120	L13
L12	L11 and L10	11	L12
L11	709/$.ccls. and ((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))	29	L11
L10	(707/1 |707/2 |707/3 |707/4 |707/5 |707/6 |707/7 |707/8 |707/9 |707/10 |707/100 |707/101 |707/102 |707/103R |707/103Y |707/103X |707/103Z |707/104.1 |707/200 |707/201 |707/202 |707/203 |707/204 |707/205 |707/206).ccls. and ((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))	152	L10
L9	((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$)).clm.	2	L9
L8	((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$)).ab.	0	L8
L7	((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$)).ti.	0	L7
L6	((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$))).clm.	1	L6
L5	((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$))).ab.	0	L5
L4	((((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$))).ti.	0	L4
L3	(((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays and (database near search$ or quer$))	183	L3
L2	((categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$))and displays	225	L2

L1	(categor$ same class or classif$ same group) and (topic same hierarch$) and (map$ same keyword$ or word$)	257	L1

END OF SEARCH HISTORY

Terms used **optimizing search by showing results** Found **146,566** of **184,245**

Sort results by: relevance
Display results: expanded form
Save results to a Binder
Search Tips
Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next
Best 200 shown Relevance scale

1 Optimizing search by showing results in context
Susan Dumais, Edward Cutrell, Hao Chen
March 2001 **Proceedings of the SIGCHI conference on Human factors in computing systems**
Publisher: ACM Press
Full text available: pdf(804.19 KB) Additional Information: full citation, abstract, references, citings, index terms

We developed and evaluated seven interfaces for integrating semantic category information with Web search results. List interfaces were based on the familiar ranked-listing of search results, sometimes augmented with a category name for each result. Category interfaces also showed page titles and/or category names, but re-organized the search results so that items in the same category were grouped together visually. Our user studies show that all Category interfaces were more effective than ...

Keywords: World Wide Web, focus-in-context, search, text categorization, usability, user interface, user study

2 Learning evaluation functions to improve optimization by local search
Justin Boyan, Andrew W. Moore
September 2001 **The Journal of Machine Learning Research**, Volume 1
Publisher: MIT Press
Full text available: pdf(643.21 KB) Additional Information: full citation, abstract, citings

This paper describes algorithms that learn to improve search performance on large-scale optimization tasks. The main algorithm, STAGE, works by learning an evaluation function that predicts the outcome of a local search algorithm, such as hillclimbing or Walksat, from features of states visited during search. The learned evaluation function is then used to bias future search trajectories toward better optima on the same problem. Another algorithm, X-STAGE, transfers previously learned evaluation ...

3 Optimizing result prefetching in web search engines with segmented indices
Ronny Lempel, Shlomo Moran
February 2004 **ACM Transactions on Internet Technology (TOIT)**, Volume 4 Issue 1
Publisher: ACM Press
Full text available: pdf(183.97 KB) Additional Information: full citation, abstract, references, index terms

We study the process in which search engines with segmented indices serve queries. In particular, we investigate the number of result pages that search engines should prepare during the query processing phase.Search engine users have been observed to browse through very few pages of results for queries that they submit. This behavior of users suggests that prefetching many results upon processing an initial query is not efficient, since most of the prefetched results will not be requested by the ...

Keywords: Distributed inverted indices, prefetching, search engines

4 Fast and efficient searches for effective optimization-phase sequences

Prasad A. Kulkarni, Stephen R. Hines, David B. Whalley, Jason D. Hiser, Jack W. Davidson, Douglas L. Jones

June 2005 **ACM Transactions on Architecture and Code Optimization (TACO)**, Volume 2 Issue 2

Publisher: ACM Press

Full text available: pdf(1.69 MB) Additional Information: full citation, abstract, references, index terms

It has long been known that a fixed ordering of optimization phases will not produce the best code for every application. One approach for addressing this phase-ordering problem is to use an evolutionary algorithm to search for a specific sequence of phases for each module or function. While such searches have been shown to produce more efficient code, the approach can be extremely slow because the application is compiled and possibly executed to evaluate each sequence's effectiveness. Consequen ...

Keywords: Phase ordering, genetic algorithms, interactive compilation

5 Fast searches for effective optimization phase sequences

Prasad Kulkarni, Stephen Hines, Jason Hiser, David Whalley, Jack Davidson, Douglas Jones

June 2004 **ACM SIGPLAN Notices , Proceedings of the ACM SIGPLAN 2004 conference on Programming language design and implementation PLDI '04**, Volume 39 Issue 6

Publisher: ACM Press

Full text available: pdf(862.40 KB) Additional Information: full citation, abstract, references, citings, index terms

It has long been known that a fixed ordering of optimization phases will not produce the best code for every application. One approach for addressing this phase ordering problem is to use an evolutionary algorithm to search for a specific sequence of phases for each module or function. While such searches have been shown to produce more efficient code, the approach can be extremely slow because the application is compiled and executed to evaluate each sequence's effectiveness. Consequently, evol ...

Keywords: genetic algorithms, interactive compilation, phase ordering

6 Compilation: In search of near-optimal optimization phase orderings

Prasad A. Kulkarni, David B. Whalley, Gary S. Tyson, Jack W. Davidson

June 2006 **Proceedings of the 2006 ACM SIGPLAN/SIGBED conference on Language, compilers and tool support for embedded systems LCTES '06**

Publisher: ACM Press

Full text available: pdf(197.28 KB) Additional Information: full citation, abstract, references, index terms

Phase ordering is a long standing challenge for traditional optimizing compilers. Varying the order of applying optimization phases to a program can produce different code, with potentially significant performance variation amongst them. A key insight to addressing the phase ordering problem is that many different optimization sequences produce the

same code. In an earlier study, we used this observation to restate the phase ordering problem to concentrate on finding all distinct function ...

Keywords: exhaustive search, genetic algorithms, phase ordering

7 Accelerating the convergence of random search methods for discrete stochastic optimization

Sigrún Andradóttir

October 1999 **ACM Transactions on Modeling and Computer Simulation (TOMACS)**, Volume 9 Issue 4

Publisher: ACM Press

Full text available: pdf(247.02 KB) Additional Information: full citation, abstract, references, citings, index terms, review

We discuss the choice of the estimation of the optimal solution when random search methods are applied to solve discrete stochastic optimization problems. At the present time, such optimization methods usually estimate the optimal solution using either the feasible solution the method is currently exploring or the feasible solution visited most often so far by the method. We propose using all the observed objective function values generated as the random search method moves around the feasi ...

Keywords: accelerated convergence, convergence rate, discrete stochastic optimization, estimating the optimal solution, random search methods

8 Variable-depth trie index optimization: theory and experimental results

R. Ramesh, A. J. G. Babu, J. Peter Kincaid

March 1989 **ACM Transactions on Database Systems (TODS)**, Volume 14 Issue 1

Publisher: ACM Press

Full text available: pdf(2.59 MB) Additional Information: full citation, abstract, references, citings, index terms, review

We develop an efficient approach to Trie index optimization. A Trie is a data structure used to index a file having a set of attributes as record identifiers. In the proposed methodology, a file is horizontally partitioned into subsets of records using a Trie index whose depth of indexing is allowed to vary. The retrieval of a record from the file proceeds by "stepping through" the index to identify a subset of records in the file in which a binary search is per ...

9 Evolutionary computation and optimization (ECO): Solving the maximum clique problem by k-opt local search

Kengo Katayama, Akihiro Hamamoto, Hiroyuki Narihisa

March 2004 **Proceedings of the 2004 ACM symposium on Applied computing**

Publisher: ACM Press

Full text available: pdf(150.12 KB) Additional Information: full citation, abstract, references

This paper presents a local search algorithm based on variable depth search, called the *k-opt local search*, for the maximum clique problem. The *k-opt* local search performs add and drop moves, each of which can be interpreted as 1-opt move, to search a *k-opt* neighborhood solution at each iteration until no better *k*-opt neighborhood solution can be found. To evaluate our *k*-opt local search algorithm, we repeatedly apply the local search for each of DIMACS benchma ...

Keywords: combinatorial optimization, local search, maximum clique problem, neighborhood, variable depth search

10 How to optimize proof-search in modal logics: new methods of proving redundancy criteria for sequent calculi

Andrei Voronkov

April 2001 **ACM Transactions on Computational Logic (TOCL)**, Volume 2 Issue 2

Publisher: ACM Press

Full text available: pdf(277.71 KB) Additional Information: full citation, abstract, references, citings, index terms

We present a bottom-up decision procedure for propositional modal logic K based on the inverse method. The procedure is based on the "inverted" version of a sequent calculus. To restrict the search space, we prove a number of redundancy criteria for derivations in the sequent calculus. We introduce a new technique of proving redundancy criteria, based on the analysis of tableau-based derivations in K. Moreover, another new technique is based on so-called

Keywords: description logics, inverse method, modal logic, proof-search, theorem proving

11 Web search and navigation: Optimizing search engines using clickthrough data

Thorsten Joachims

July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**

Publisher: ACM Press

Full text available: pdf(954.36 KB) Additional Information: full citation, abstract, references, citings, index terms

This paper presents an approach to automatically optimizing the retrieval quality of search engines using clickthrough data. Intuitively, a good information retrieval system should present relevant documents high in the ranking, with less relevant documents following below. While previous approaches to learning retrieval functions from examples exist, they typically require training data generated from relevance judgments by experts. This makes them difficult and expensive to apply. The goal of ...

12 Ranking: Boolean + ranking: querying a database by k-constrained optimization

Zhen Zhang, Seung-won Hwang, Kevin Chen-Chuan Chang, Min Wang, Christian A. Lang, Yuan-chi Chang

June 2006 **Proceedings of the 2006 ACM SIGMOD international conference on Management of data SIGMOD '06**

Publisher: ACM Press

Full text available: pdf(316.14 KB) Additional Information: full citation, abstract, references, index terms

The wide spread of databases for managing structured data, compounded with the expanded reach of the Internet, has brought forward interesting *data retrieval* and *analysis* scenarios to RDBMS. In such settings, queries often take the form of *k-constrained optimization*, with a Boolean constraint and a numeric optimization expression as the goal function, retrieving only the top-*k* tuples. This paper proposes the concept of supporting such queries, as their nature i ...

Keywords: A* search, constrained optimization, index, query processing, top-k query

13 Session 7B: Approximate local search in combinatorial optimization

James B. Orlin, Abraham P. Punnen, Andreas S. Schulz

January 2004 **Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms**

Publisher: Society for Industrial and Applied Mathematics

Full text available: pdf(200.80 KB) Additional Information: full citation, abstract, references, citings

Local search algorithms for combinatorial optimization problems are in general of pseudopolynomial running time and polynomial-time algorithms are often not known for finding locally optimal solutions for NP-hard optimization problems. We introduce the concept of ε-local optimality and show that an ε-local optimum can be identified in time polynomial in the problem size and $1=\varepsilon$ whenever the corresponding neighborhood can be searched in polynomial time, for $\varepsilon > 0$. If the ne ...

14 Combining Models and Guided Empirical Search to Optimize for Multiple Levels of the Memory Hierarchy

Chun Chen, Jacqueline Chame, Mary Hall

March 2005 **Proceedings of the international symposium on Code generation and optimization CGO '05**

Publisher: IEEE Computer Society

Full text available: pdf(330.21 KB) Additional Information: full citation, abstract, index terms

This paper describes an algorithm for simultaneously optimizing across multiple levels of the memory hierarchy for dense-matrix computations. Our approach combines compiler models and heuristics with guided empirical search to take advantage of their complementary strengths. The models and heuristics limit the search to a small number of candidate implementations, and the empirical results provide the most accurate information to the compiler to select among candidates and tune optimization para ...

15 Mismatch analysis and direct yield optimization by specwise linearization and feasibility-guided search

Frank Schenkel, Michael Pronath, Stephen Zizala, Robert Schwencker, Helmut Graeb, Kurt Antreich

June 2001 **Proceedings of the 38th conference on Design automation**

Publisher: ACM Press

Full text available: pdf(187.40 KB) Additional Information: full citation, abstract, references, citings, index terms

We present a new method for mismatch analysis and automatic yield optimization of analog integrated circuits with respect to global, local and operational tolerances. Effectiveness and efficiency of yield estimation and optimization are guaranteed by consideration of feasibility regions and by performance linearization at worst-case points. The proposed methods were successfully applied to two example circuits for an industrial fabrication process.

16 Evolutionary combinatorial optimization: papers: Maximum cardinality matchings on trees by randomized local search

Oliver Giel, Ingo Wegenerraise

July 2006 **Proceedings of the 8th annual conference on Genetic and evolutionary computation GECCO '06**

Publisher: ACM Press

Full text available: pdf(221.43 KB) Additional Information: full citation, abstract, references, index terms

To understand the working principles of randomized search heuristics like evolutionary algorithms they are analyzed on optimization problems whose structure is well-studied. The idea is to investigate when it is possible to simulate clever optimization techniques for combinatorial optimization problems by random search. The maximum matching problem is well suited for this approach since long augmenting paths do not allow immediate improvements by local changes. It is known that randomized search ...

Keywords: evolutionary algorithms, maximum cardinality matchings, randomized local search, runtime analysis

17 Speedup learning for repair-based search by identifying redundant steps
Shaul Markovitch, Asaf Shatil
December 2003 **The Journal of Machine Learning Research**, Volume 4
Publisher: MIT Press
Full text available: pdf(257.03 KB) Additional Information: full citation, abstract, references, index terms

Repair-based search algorithms start with an initial solution and attempt to improve it by iteratively applying repair operators. Such algorithms can often handle large-scale problems that may be difficult for systematic search algorithms. Nevertheless, the computational cost of solving such problems is still very high. We observed that many of the repair steps applied by such algorithms are redundant in the sense that they do not eventually contribute to finding a solution. Such redundant steps ...

18 A recursive random search algorithm for network parameter optimization
Tao Ye, Shivkumar Kalyanaraman
December 2004 **ACM SIGMETRICS Performance Evaluation Review**, Volume 32 Issue 3
Publisher: ACM Press
Full text available: pdf(916.09 KB) Additional Information: full citation, abstract, references

This paper proposes a new heuristic search algorithm, Recursive Random Search(RRS), for black-box optimization problems. Specifically, this algorithm is designed for the dynamical parameter optimization of network protocols which emphasizes on obtaining good solutions within a limited time frame rather than full optimization. The RRS algorithm is based on the initial high-efficiency property of random sampling and attempts to maintain this high-efficiency by constantly "restarting" random sampli ...

19 Genetic algorithms: A comparison study between genetic algorithms and bayesian optimize algorithms by novel indices
Naoki Mori, Masayuki Takeda, Keinosuke Matsumoto
June 2005 **Proceedings of the 2005 conference on Genetic and evolutionary computation GECCO '05**
Publisher: ACM Press
Full text available: pdf(1.86 MB) Additional Information: full citation, abstract, references, index terms

Genetic Algorithms (GAs) are a search and optimization technique based on the mechanism of evolution. Recently, another sort of population-based optimization method called Estimation of Distribution Algorithms (EDAs) have been proposed to solve the GA's defects. Although several comparison studies between GAs and EDAs have been made, little is known about differences of statistical features between them. In this paper, we propose new statistical indices which are based on the concepts of crossov ...

Keywords: bayesian optimization algorithms, diversity, genetic algorithms, population-based optimization methods

20 Variable-sample methods for stochastic optimization
Tito Homem-De-Mello
April 2003 **ACM Transactions on Modeling and Computer Simulation (TOMACS)**, Volume 13 Issue 2
Publisher: ACM Press
Full text available: pdf(244.36 KB) Additional Information: full citation, abstract, references, index terms

In this article we discuss the application of a certain class of Monte Carlo methods to stochastic optimization problems. Particularly, we study *variable-sample* techniques, in which the objective function is replaced, *at each iteration*, by a sample average

approximation. We first provide general results on the *schedule* of sample sizes, under which variable-sample methods yield consistent estimators as well as bounds on the estimation error. Because the convergence analysis i ...

Keywords: Monte Carlo methods, pathwise bounds, random search, stochastic optimization

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

Results 1 - 20 of 200 Result page: **1** 2 3 4 5 6 7 8 9 10 next

Best 200 shown Relevance scale

1 Search engineering 2: Mining models of human activities from the web

Mike Perkowitz, Matthai Philipose, Kenneth Fishkin, Donald J. Patterson

May 2004 **Proceedings of the 13th international conference on World Wide Web**

Publisher: ACM Press

Full text available: pdf(582.93 KB) Additional Information: full citation, abstract, references, index terms

The ability to determine what day-to-day activity (such as cooking pasta, taking a pill, or watching a video) a person is performing is of interest in many application domains. A system that can do this requires models of the activities of interest, but model construction does not scale well: humans must specify low-level details, such as segmentation and feature selection o sensor data, and high-level structure, such as spatio-temporal relations between states of the model, for each and every ...

Keywords: activity inference, activity models, rfid, web mining

2 BANANAS: an evolutionary framework for explicit and multipath routing in the internet

H. Tahilramani Kaur, S. Kalyanaraman, A. Weiss, S. Kanwar, A. Gandhi

August 2003 **ACM SIGCOMM Computer Communication Review , Proceedings of the ACM SIGCOMM workshop on Future directions in network architecture FDNA '03,** Volume 33 Issue 4

Publisher: ACM Press

Full text available: pdf(585.15 KB) Additional Information: full citation, abstract, references, citings

Today the Internet offers a single path between end-systems even though it intrinsically has a large multiplicity of paths. This paper proposes an evolutionary architectural framework "BANANAS" aimed at simplifying the introduction of multipath routing in the Internet. The framework starts with the observation that a path can be encoded as a short hash ("PathID") of a sequence of globally known identifiers. The PathID therefore has global significance (unlike MPLS or ATM labels). This property a ...

3 Natural-language retrieval of images based on descriptive captions

Eugene J. Guglielmo, Neil C. Rowe

July 1996 **ACM Transactions on Information Systems (TOIS),** Volume 14 Issue 3

Publisher: ACM Press

Full text available: pdf(572.05 KB) Additional Information: full citation, abstract, references, citings, index terms, review

We describe a prototype intelligent information retrieval system that uses natural-language understanding to efficiently locate captioned data. Multimedia data generally require captions tc explain their features and significance. Such descriptive captions often rely on long nominal compounds (strings of consecutive nouns) which create problems of disambiguating word sence In our system, captions and user queries are parsed and interpreted to produce a logical form using a detailed theory ...

Keywords: captions, multimedia database, type hierarchy

4 Data mining (DM): Optimizing subset queries: a step towards SQL-based inductive databases for itemsets

Cyrille Masson, Céline Robardet, Jean-François Boulicaut
March 2004 **Proceedings of the 2004 ACM symposium on Applied computing**
Publisher: ACM Press
Full text available: pdf(170.27 KB) Additional Information: full citation, abstract, references

Storing sets and querying them (e.g., subset queries that provide all supersets of a given set) i: known to be difficult within relational databases. We consider that being able to query efficientl both transactional data and materialized collections of sets by means of standard query language is an important step towards practical inductive databases. Indeed, data mining query languages like MINE RULE extract collections of association rules whose components are sets into relational tables. Po ...

Keywords: inductive databases, itemset post-processing

5 Optimal signature extraction and information loss

Christos Faloutsos, Stavros Christodoulakis
September 1987 **ACM Transactions on Database Systems (TODS)**, Volume 12 Issue 3
Publisher: ACM Press
Full text available: pdf(1.98 MB) Additional Information: full citation, abstract, references, citings, index terms

Signature files seem to be a promising access method for text and attributes. According to this method, the documents (or records) are stored sequentially in one file ("text file"), while abstractions of the documents ("signatures") are stored sequentially in another file ("signature file"). In order to resolve a query, the signature file is scanned first, and many nonqualifying documents are immediately rejected. We develop a framework that includes primary key hashing, multiattribute hash ...

6 Incomplete information costs and database design

Haim Mendelson, Aditya N. Saharia
June 1986 **ACM Transactions on Database Systems (TODS)**, Volume 11 Issue 2
Publisher: ACM Press
Full text available: pdf(1.35 MB) Additional Information: full citation, abstract, references, citings, index terms

This paper presents a methodology for trading-off the cost of incomplete information against th data-related costs in the design of database systems. It investigates how the usage patterns of the database, defined by the characteristics of information requests presented to it, affect its conceptual design. The construction of minimum-cost answers to information requests for a variety of query types and cost structures is also studied. The resulting costs of incomplete database information ...

7 Database principles: Logic-based web information extraction

Georg Gottlob, Christoph Koch
June 2004 **ACM SIGMOD Record**, Volume 33 Issue 2

Publisher: ACM Press
Full text available: pdf(1.74 MB) Additional Information: full citation, abstract, references

The Web wrapping proble, i.e., the problem of extracting structured information from HTML documents, is one of great practical importance. The often observed *information overload* that users of the Web experience witnesses the lack of intelligent and encompassing Web services that provide high-quality collected and value-added inforamtion. The Web wrapping problem ha been addressed by a significant amount of research work. Previous work can be classified into two categories, depending on ...

8 Final report of the ANSI/X3/SPARC DBS-SG relational database task group
July 1982 **ACM SIGMOD Record**, Volume 12 Issue 4
Publisher: ACM Press
Full text available: pdf(4.69 MB) Additional Information: full citation

9 Surrogate subsets: a free space management strategy for the index of a text retrieval system
F. J. Burkowski
December 1989 **Proceedings of the 13th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: ACM Press
Full text available: pdf(1.44 MB) Additional Information: full citation, abstract, references, citings, index terms

This paper presents a new data structure and an associated strategy to be utilized by indexing facilities for text retrieval systems. The paper starts by reviewing some of the goals that may be considered when designing such an index and continues with a small survey of various current strategies. It then presents an indexing strategy referred to as surrogate subsets discussing its appropriateness in the light of the specified goals. Various design issues and implementation details are disc ...

10 Database theory, technology, and applications (DTTA): Fingerprinting relational databases
Fei Guo, Jianmin Wang, Deyi Li
April 2006 **Proceedings of the 2006 ACM symposium on Applied computing SAC '06**
Publisher: ACM Press
Full text available: pdf(165.18 KB) Additional Information: full citation, abstract, references, index terms

In this paper, we propose a fingerprinting solution to protect valuable numeric relational data from illegal duplications and redistributions. We introduce a twice-embedding scheme. In the first embedding process, we embed a unique fingerprint to identify each recipient to whom the relational data is distributed. The embedding process is controlled by a secret key. Meanwhile, the fingerprint can be detected using the same secret key to prove ownership at a numerical confidence level. The second ...

Keywords: database, fingerprint, robustness, watermark

11 QProber: A system for automatic classification of hidden-Web databases
Luis Gravano, Panagiotis G. Ipeirotis, Mehran Sahami
January 2003 **ACM Transactions on Information Systems (TOIS)**, Volume 21 Issue 1
Publisher: ACM Press
Full text available: pdf(3.62 MB) Additional Information: full citation, abstract, references, citings, index terms

The contents of many valuable Web-accessible databases are only available through search interfaces and are hence invisible to traditional Web "crawlers." Recently, commercial Web sites have started to manually organize Web-accessible databases into Yahoo!-like hierarchical

classification schemes. Here we introduce QProber, a modular system that automates this classification process by using a small number of query probes, generated by document classifiers. QProber can use a variety of types of ...

Keywords: Database classification, Web databases, hidden Web

12 PR-Miner: automatically extracting implicit programming rules and detecting violations in large software code

Zhenmin Li, Yuanyuan Zhou

September 2005 **ACM SIGSOFT Software Engineering Notes , Proceedings of the 10th European software engineering conference held jointly with 13th ACM SIGSOFT international symposium on Foundations of software engineering ESEC/FSE-13**, Volume 30 Issue 5

Publisher: ACM Press

Full text available: pdf(228.58 KB) Additional Information: full citation, abstract, references, index terms

Programs usually follow many *implicit* programming rules, most of which are too tedious to be documented by programmers. When these rules are violated by programmers who are unaware of or forget about them, defects can be easily introduced. Therefore, it is highly desirable to have tools to automatically extract such rules and also to automatically detect violations. Previous work in this direction focuses on simple function-pair based programming rules and additionally requires programmer ...

Keywords: automated specification generation, automated violation detection, data mining for software engineering, pattern recognition, programming rules, static analysis

13 Session 11: multimedia analysis and retrieval: VQ-index: an index structure for similarity searching in multimedia databases

Ertem Tuncel, Hakan Ferhatosmanoglu, Kenneth Rose

December 2002 **Proceedings of the tenth ACM international conference on Multimedia**

Publisher: ACM Press

Full text available: pdf(525.17 KB) Additional Information: full citation, abstract, references, citings

In this paper, we introduce a novel indexing technique based on efficient compression of the feature space for approximate similarity searching in large multimedia databases. Its main novelty is that state-of-the-art tools from the discipline of data compression are adopted to optimize the complexity-performance tradeoff in large data sets. The design procedure optimize the query access time by jointly accounting for both database distribution and query statistics. We achieve efficient compress ...

Keywords: approximate similarity searching, clustering, indexing, retrieved information reduction, retrieved set reduction, vector quantization

14 Performance analysis in the software lifecycle: The Sisyphus database retrieval software performance antipattern

Robert F. Dugan, Ephraim P. Glinert, Ali Shokoufandeh

July 2002 **Proceedings of the 3rd international workshop on Software and performance WOSP '02**

Publisher: ACM Press

Full text available: pdf(1.18 MB) Additional Information: full citation, abstract, references, citings, index terms

In this paper we propose the Sisyphus database retrieval software performance antipattern. The antipattern occurs in application designs that process large, frequently accessed lists stored in a relational database, but display only a small subset to the user. Software Performance

Engineering (SPE) techniques are used to analyze the antipattern. Four solutions are evaluated: rownum and index, upper/lower bound, sequence numbering, and caching. We discuss the real world challenges of correcting t ...

Keywords: antipatterns, patterns

15 Essential classification rule sets

Elena Baralis, Silvia Chiusano
December 2004 **ACM Transactions on Database Systems (TODS)**, Volume 29 Issue 4
Publisher: ACM Press
Full text available: pdf(479.09 KB) Additional Information: full citation, abstract, references, index terms

Given a class model built from a dataset including labeled data, classification assigns a new dat object to the appropriate class. In associative classification the class model (i.e., the classifier) i a set of association rules. Associative classification is a promising technique for the generation c highly accurate classifiers. In this article, we present a compact form which encodes without information loss the classification knowledge available in a classification rule set. This form incl ...

Keywords: Association rules, associative classification, concise representations

16 Monadic datalog and the expressive power of languages for Web information extraction

Georg Gottlob, Christoph Koch
January 2004 **Journal of the ACM (JACM)**, Volume 51 Issue 1
Publisher: ACM Press
Full text available: pdf(277.94 KB) Additional Information: full citation, abstract, references, citings, index terms

Research on information extraction from Web pages (wrapping) has seen much activity recently (particularly systems implementations), but little work has been done on formally studying the expressiveness of the formalisms proposed or on the theoretical foundations of wrapping. In thi paper, we first study monadic datalog over trees as a wrapping language. We show that this simple language is equivalent to monadic second order logic (MSO) in its ability to specify wrappers. We believe that MSO ha ...

Keywords: Complexity, HTML, MSO, expressiveness, information extraction, monadic datalog, regular tree languages, web wrapping

17 Extracting predicates from mining models for efficient query evaluation

Surajit Chaudhuri, Vivek Narasayya, Sunita Sarawagi
September 2004 **ACM Transactions on Database Systems (TODS)**, Volume 29 Issue 3
Publisher: ACM Press
Full text available: pdf(698.37 KB) Additional Information: full citation, abstract, references, index terms

Modern relational database systems are beginning to support ad hoc queries on mining models. In this article, we explore novel techniques for optimizing queries that contain predicates on the results of application of mining models to relational data. For such queries, we use the internal structure of the mining model to automatically derive traditional database predicates. We present algorithms for deriving such predicates for a large class of popular discrete mining models: decision trees, nai ...

Keywords: Complex predicate optimization, simpler rules from complex predictive functions

18 XTRACT: a system for extracting document type descriptors from XML documents

Minos Garofalakis, Aristides Gionis, Rajeev Rastogi, S. Seshadri, Kyuseok Shim
May 2000 **ACM SIGMOD Record , Proceedings of the 2000 ACM SIGMOD international conference on Management of data SIGMOD '00**, Volume 29 Issue 2
Publisher: ACM Press
Full text available: pdf(209.66 KB) Additional Information: full citation, abstract, references, citings, index terms

XML is rapidly emerging as the new standard for data representation and exchange on the Web. An XML document can be accompanied by a *Document Type Descriptor* (DTD) which plays the role of a schema for an XML data collection. DTDs contain valuable information on the structure of documents and thus have a crucial role in the efficient storage of XML data, as well as the effective formulation and optimization of XML queries. In this paper, we propose XTRACT, a novel system for inferring a ...

19 The deductive synthesis of database transactions

Xiaolei Qian
December 1993 **ACM Transactions on Database Systems (TODS)**, Volume 18 Issue 4
Publisher: ACM Press
Full text available: pdf(3.27 MB) Additional Information: full citation, references, index terms

Keywords: database programming, deductive tableau, integrity constraints, search control, transaction logic, transaction synthesis

20 A system for discovering relationships by feature extraction from text databases

Jack G. Conrad, Mary Hunter Utt
August 1994 **Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval**
Publisher: Springer-Verlag New York, Inc.
Full text available: pdf(911.76 KB) Additional Information: full citation, references, citings, index terms, review